MVO: Origins and Current Problems

chris bemis
Whitebox Advisors

September 11, 2009

Modern portfolio theory establishes

- A necessity for a forecasted reward
- ...and a specified risk over this same period.
- These two quantities are balanced against each other in some 'optimal' way.

Modern portfolio theory establishes

- A necessity for a forecasted reward
- ...and a specified risk over this same period.
- These two quantities are balanced against each other in some 'optimal' way.

Modern portfolio theory establishes

- A necessity for a forecasted reward
- ...and a specified risk over this same period.
- These two quantities are balanced against each other in some 'optimal' way.

Modern portfolio theory establishes

- A necessity for a forecasted reward
- ...and a specified risk over this same period.
- These two quantities are balanced against each other in some 'optimal' way.

Modern portfolio theory establishes

- A necessity for a forecasted reward
- ...and a specified risk over this same period.
- These two quantities are balanced against each other in some ’optimal’ way.

In the final analysis, the vagueness of the above is why portfolio optimization is a tool rather than a solution.

Markowitz in 1952 established what is likely the first portfolio optimization problem; namely

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

A QP problem with linear constraints.
It was a significant achievement, and it has been widely adopted. Enough so that we know it doesn't work.

Markowitz in 1952 established what is likely the first portfolio optimization problem; namely

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

A QP problem with linear constraints.

It was a significant achievement, and it has been widely adopted. Enough so that we know it doesn't work.

Markowitz in 1952 established what is likely the first portfolio optimization problem; namely

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

A QP problem with linear constraints.
It was a significant achievement, and it has been widely adopted. Enough so that we know it doesn't work.

Markowitz in 1952 established what is likely the first portfolio optimization problem; namely

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

A QP problem with linear constraints.
It was a significant achievement, and it has been widely adopted. Enough so that we know it doesn't work.

Markowitz in 1952 established what is likely the first portfolio optimization problem; namely

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

A QP problem with linear constraints.
It was a significant achievement, and it has been widely adopted. Enough so that we know it doesn't work.

The above is generally referred to as mean-variance optimization.
In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

The above is generally referred to as mean-variance optimization.

In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

The above is generally referred to as mean-variance optimization.
In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

The above is generally referred to as mean-variance optimization.
In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

The above is generally referred to as mean-variance optimization.

In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

This formulation is normative and not empirical.

The above is generally referred to as mean-variance optimization.

In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

This formulation is normative and not empirical.

Financial data are nonstationary.

The above is generally referred to as mean-variance optimization.

In terms of the original discussion, we are equating

- Reward with expected return
- Risk with portfolio variance
- We are minimizing risk subject to a minimum expected return

This formulation is normative and not empirical.

Financial data are nonstationary.

Additionally, measurements have error.

In time series, we often think of observations through time as random variables.

Generally, in time series, we want to have the joint distribution of time lagged variables to only depend on the lag - not the time.

For example, we might expect that the correlation between two stocks' returns might be the same over distinct 200 day periods.

Or, a covariance matrix might be assumed to remain the same in the future.

Stopping at the second moment is an example of weak stationarity.

In time series, we often think of observations through time as random variables.

Generally, in time series, we want to have the joint distribution of time lagged variables to only depend on the lag - not the time.

For example, we might expect that the correlation between two stocks' returns might be the same over distinct 200 day periods.

Or, a covariance matrix might be assumed to remain the same in the future.

Stopping at the second moment is an example of weak stationarity.

In time series, we often think of observations through time as random variables.

Generally, in time series, we want to have the joint distribution of time lagged variables to only depend on the lag - not the time.

For example, we might expect that the correlation between two stocks' returns might be the same over distinct 200 day periods.

Or, a covariance matrix might be assumed to remain the same in the future.

Stopping at the second moment is an example of weak stationarity.

In time series, we often think of observations through time as random variables.

Generally, in time series, we want to have the joint distribution of time lagged variables to only depend on the lag - not the time.

For example, we might expect that the correlation between two stocks' returns might be the same over distinct 200 day periods. Or, a covariance matrix might be assumed to remain the same in the future.

Stopping at the second moment is an example of weak stationarity.

In time series, we often think of observations through time as random variables.
Generally, in time series, we want to have the joint distribution of time lagged variables to only depend on the lag - not the time.
For example, we might expect that the correlation between two stocks' returns might be the same over distinct 200 day periods.
Or, a covariance matrix might be assumed to remain the same in the future.
Stopping at the second moment is an example of weak stationarity.

The question arises...

Can we test if financial time series are weakly stationary?

We don't have to look very far:

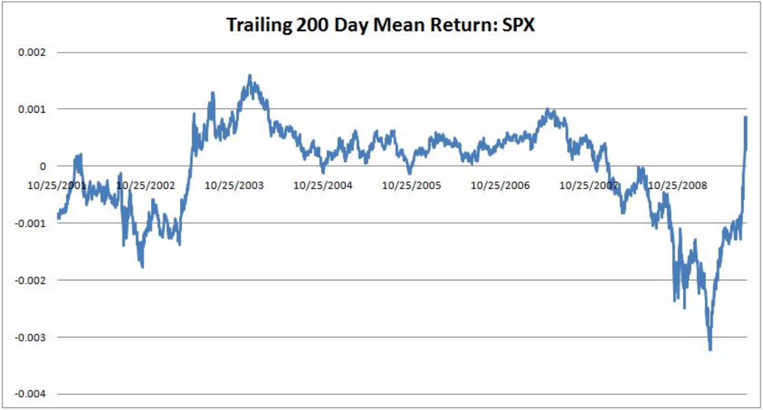
The question arises...

Can we test if financial time series are weakly stationary?

We don't have to look very far:

The question arises...
Can we test if financial time series are weakly stationary?
We don't have to look very far:

The question arises...
Can we test if financial time series are weakly stationary?
We don't have to look very far:



Putting aside practical questions for a moment, there are interesting features of the problem

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

like, what happens when we solve the problem for various α's?

Putting aside practical questions for a moment, there are interesting features of the problem

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

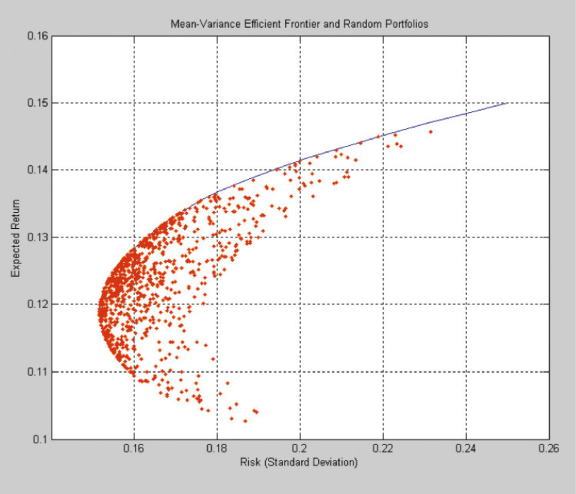
like, what happens when we solve the problem for various α's?

Putting aside practical questions for a moment, there are interesting features of the problem

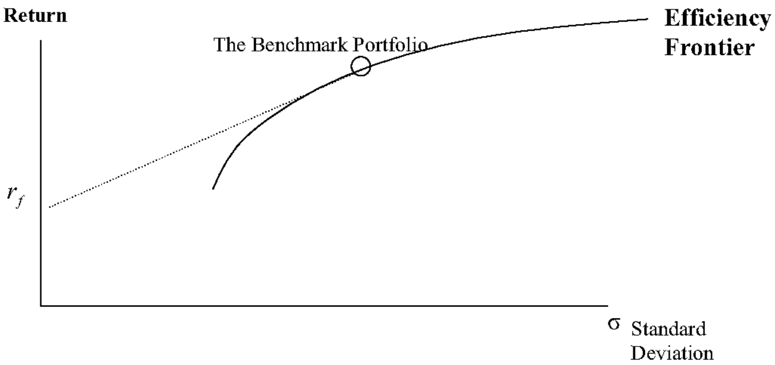
$$\begin{aligned} \text{minimize} \quad & \mathbf{w}'\Sigma\mathbf{w} \\ \text{subject to} \quad & \mu'\mathbf{w} \geqslant \alpha \\ & \mathbf{1}'\mathbf{w} = 1 \\ & \mathbf{w} \geqslant 0, \end{aligned}$$

like, what happens when we solve the problem for various α's?

We get something called the efficient frontier.

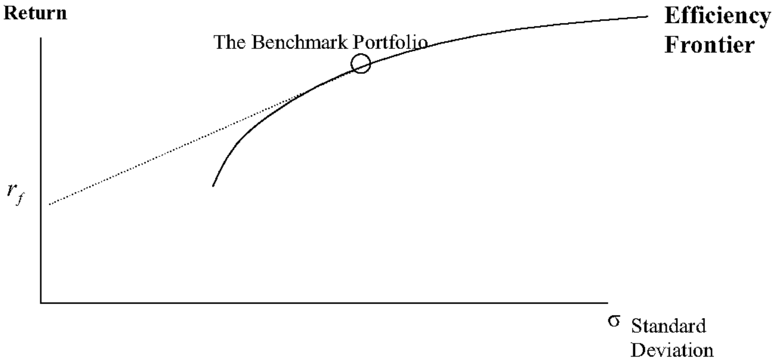


Now, what if there is a risk free asset. One with no volatility. We get the 'market portfolio' and β



The slope of the market line is the so-called Sharpe Ratio

Now, what if there is a risk free asset. One with no volatility. We get the 'market portfolio' and β



The slope of the market line is the so-called Sharpe Ratio

If we put on our normative hats, we start making claims:

- You'd always want a portfolio *on* the efficient frontier
- If there is a risk free asset, you'd always want to be on the market line
- Ergo, all risk is systemic and you just need to figure out how much of the market portfolio you want
- huh? Well,

$$r_i - r_f = \beta(r_m - r_f)$$

If we put on our normative hats, we start making claims:

- You'd always want a portfolio *on* the efficient frontier
- If there is a risk free asset, you'd always want to be on the market line
- Ergo, all risk is systemic and you just need to figure out how much of the market portfolio you want
- huh? Well,

$$r_i - r_f = \beta(r_m - r_f)$$

If we put on our normative hats, we start making claims:

- You'd always want a portfolio *on* the efficient frontier
- If there is a risk free asset, you'd always want to be on the market line
- Ergo, all risk is systemic and you just need to figure out how much of the market portfolio you want
- huh? Well,

$$r_i - r_f = \beta(r_m - r_f)$$

If we put on our normative hats, we start making claims:

- You'd always want a portfolio *on* the efficient frontier
- If there is a risk free asset, you'd always want to be on the market line
- Ergo, all risk is systemic and you just need to figure out how much of the market portfolio you want
- huh? Well,

$$r_i - r_f = \beta(r_m - r_f)$$

If we put on our normative hats, we start making claims:

- You'd always want a portfolio *on* the efficient frontier
- If there is a risk free asset, you'd always want to be on the market line
- Ergo, all risk is systemic and you just need to figure out how much of the market portfolio you want
- huh? Well,

$$r_i - r_f = \beta(r_m - r_f)$$

We also obtain another optimization problem, namely, finding the optimal Sharpe ratio:

$$\max_{\{w \in \mathcal{W}\}} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}$$

Here we are choosing a specific point on the efficient frontier without specifying a minimum return

We also obtain another optimization problem, namely, finding the optimal Sharpe ratio:

$$\max_{\{w \in \mathcal{W}\}} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}$$

Here we are choosing a specific point on the efficient frontier without specifying a minimum return

We also obtain another optimization problem, namely, finding the optimal Sharpe ratio:

$$\max_{\{w \in \mathcal{W}\}} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}$$

Here we are choosing a specific point on the efficient frontier without specifying a minimum return

The results are neat.

You have to ask yourself, though, what made such an oversimplified statement possible?

Everything in the setting above, the Capital Asset Pricing Model, assumes that means are stable through time. And so are covariances.

The good news for the future practitioner: they're not. At all.

The results are neat.

You have to ask yourself, though, what made such an oversimplified statement possible?

Everything in the setting above, the Capital Asset Pricing Model, assumes that means are stable through time. And so are covariances.

The good news for the future practitioner: they're not. At all.

The results are neat.

You have to ask yourself, though, what made such an oversimplified statement possible?

Everything in the setting above, the Capital Asset Pricing Model, assumes that means are stable through time. And so are covariances.

The good news for the future practitioner: they're not. At all.

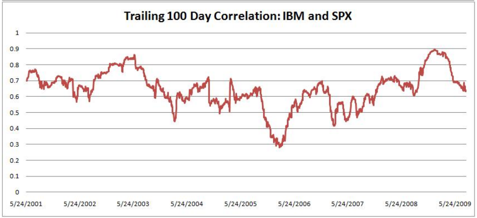
The results are neat.

You have to ask yourself, though, what made such an oversimplified statement possible?

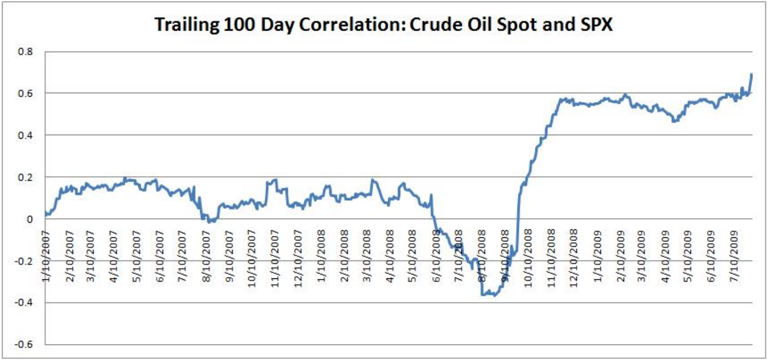
Everything in the setting above, the Capital Asset Pricing Model, assumes that means are stable through time. And so are covariances.

The good news for the future practitioner: they're not. At all.

A disclaimer: this view on the origins of modern portfolio theory may not be the norm. In fact, Markowitz was quoted as saying "Diversifying sufficiently among uncorrelated risks can reduce portfolio risk toward zero" as recently as 2008. However,



And in case that doesn't seem fair:



There are times when "correlation goes to one." Hence true diversification may fail exactly when you need it most.

We see the need to mitigate definable risks according to our understanding (model).

We see here the inherit importance of understanding our underlying; e.g., beginning with the premise that our data is nonstationary, or understanding what influences stock prices, say.

This is an empirical exercise and not a normative one.

There are times when "correlation goes to one." Hence true diversification may fail exactly when you need it most.

We see the need to mitigate definable risks according to our understanding (model).

We see here the inherit importance of understanding our underlying; e.g., beginning with the premise that our data is nonstationary, or understanding what influences stock prices, say.

This is an empirical exercise and not a normative one.

There are times when "correlation goes to one." Hence true diversification may fail exactly when you need it most.

We see the need to mitigate definable risks according to our understanding (model).

We see here the inherit importance of understanding our underlying; e.g., beginning with the premise that our data is nonstationary, or understanding what influences stock prices, say.

This is an empirical exercise and not a normative one.

There are times when ”correlation goes to one.” Hence true diversification may fail exactly when you need it most.

We see the need to mitigate definable risks according to our understanding (model).

We see here the inherit importance of understanding our underlying; e.g., beginning with the premise that our data is nonstationary, or understanding what influences stock prices, say.

This is an empirical exercise and not a normative one.

There is no right way to ‘mitigate definable risks according to our understanding’.

Again, there are tools, not answers.

Tutuncu suggests a robust optimization procedure using interval sets around expectation and the elements of the covariance matrix.

Goldfarb and Iyengar suggest that you define a linear factor model and use elliptical uncertainty sets based on regression coefficients.

There is no right way to 'mitigate definable risks according to our understanding'.

Again, there are tools, not answers.

Tutuncu suggests a robust optimization procedure using interval sets around expectation and the elements of the covariance matrix.

Goldfarb and Iyengar suggest that you define a linear factor model and use elliptical uncertainty sets based on regression coefficients.

There is no right way to 'mitigate definable risks according to our understanding'.

Again, there are tools, not answers.

Tutuncu suggests a robust optimization procedure using interval sets around expectation and the elements of the covariance matrix.

Goldfarb and Iyengar suggest that you define a linear factor model and use elliptical uncertainty sets based on regression coefficients.

There is no right way to ‘mitigate definable risks according to our understanding’.

Again, there are tools, not answers.

Tutuncu suggests a robust optimization procedure using interval sets around expectation and the elements of the covariance matrix.

Goldfarb and Iyengar suggest that you define a linear factor model and use elliptical uncertainty sets based on regression coefficients.

Both of these options have their place.

In the case of Tutuncu, the practitioner has some added flexibility, but may lose structure.

The negation of this statement is true for Goldfarb and Iyengar.

Below, we create a hybrid of the two, using a practitioner informed elliptical uncertainty set based on a linear factor model.

We are therefore motivated to construct *some* linear factor model for returns.

Both of these options have their place.

In the case of Tutuncu, the practitioner has some added flexibility, but may lose structure.

The negation of this statement is true for Goldfarb and Iyengar.

Below, we create a hybrid of the two, using a practitioner informed elliptical uncertainty set based on a linear factor model.

We are therefore motivated to construct *some* linear factor model for returns.

Both of these options have their place.

In the case of Tutuncu, the practitioner has some added flexibility, but may lose structure.

The negation of this statement is true for Goldfarb and Iyengar.

Below, we create a hybrid of the two, using a practitioner informed elliptical uncertainty set based on a linear factor model.

We are therefore motivated to construct *some* linear factor model for returns.

Both of these options have their place.

In the case of Tutuncu, the practitioner has some added flexibility, but may lose structure.

The negation of this statement is true for Goldfarb and Iyengar.

Below, we create a hybrid of the two, using a practitioner informed elliptical uncertainty set based on a linear factor model.

We are therefore motivated to construct *some* linear factor model for returns.

Both of these options have their place.
In the case of Tutuncu, the practitioner has some added flexibility, but may lose structure.
The negation of this statement is true for Goldfarb and Iyengar.
Below, we create a hybrid of the two, using a practitioner informed elliptical uncertainty set based on a linear factor model.
We are therefore motivated to construct *some* linear factor model for returns.

This linear factor model is an example of attempting to 'understand our underlying'.

One way was established by the work of Fama and French They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

This linear factor model is an example of attempting to ‘understand our underlying’.

One way was established by the work of Fama and French

They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

This linear factor model is an example of attempting to 'understand our underlying'.
One way was established by the work of Fama and French
They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

This linear factor model is an example of attempting to 'understand our underlying'.
One way was established by the work of Fama and French
They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

This linear factor model is an example of attempting to 'understand our underlying'.
One way was established by the work of Fama and French
They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

This linear factor model is an example of attempting to 'understand our underlying'.
One way was established by the work of Fama and French
They suggested that risk in equities and bonds could be explained in (large) part by

- Size
- Market Sensitivity; viz., β above
- Book to Price Ratio; i.e., a measure of value.

Fama and French do not make ex-ante forecasts.
They are explaining the past, not predicting the future.
They are proponents of the Efficient Market Hypothesis, which is capitalized so it must be true.

Fama and French do not make ex-ante forecasts.
They are explaining the past, not predicting the future.
They are proponents of the Efficient Market Hypothesis, which is capitalized so it must be true.

Fama and French do not make ex-ante forecasts.
They are explaining the past, not predicting the future.
They are proponents of the Efficient Market Hypothesis, which is capitalized so it must be true.

Fama and French do not make ex-ante forecasts.
They are explaining the past, not predicting the future.
They are proponents of the Efficient Market Hypothesis, which is capitalized so it must be true.

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

Let's ignore the EMH for a moment, and look at a cross-sectional regression model with factors:

- Size: market cap
- Value: book to price
- Leverage: debt to equity
- Asset growth: measured year over year
- Momentum and mean reversion: one month trailing returns
- We also include an intercept term

What have we done by specifying this model?

We have assumed that future returns are explained by these factor exposures, and further, that our uncertainty in our estimate is necessarily a function of our regression procedure. We have not addressed nonstationarity of factor returns, but have sought to address the nonstationarity of underlying returns.

What have we done by specifying this model?

We have assumed that future returns are explained by these factor exposures, and further, that our uncertainty in our estimate is necessarily a function of our regression procedure. We have not addressed nonstationarity of factor returns, but have sought to address the nonstationarity of underlying returns.

What have we done by specifying this model?

We have assumed that future returns are explained by these factor exposures, and further, that our uncertainty in our estimate is necessarily a function of our regression procedure.

We have not addressed nonstationarity of factor returns, but have sought to address the nonstationarity of underlying returns.

What have we done by specifying this model?

We have assumed that future returns are explained by these factor exposures, and further, that our uncertainty in our estimate is necessarily a function of our regression procedure.

We have not addressed nonstationarity of factor returns, but have sought to address the nonstationarity of underlying returns.

If r is a forecast, and V is a design matrix, we have

$$r = V^T \cdot f + \epsilon$$

The vector f is determined here by regression, and so is the assumed distribution of ϵ.
Below we assume $\epsilon \sim N(0, D)$, with D diagonal.

If r is a forecast, and V is a design matrix, we have

$$r = V^T \cdot f + \epsilon$$

The vector f is determined here by regression, and so is the assumed distribution of ϵ.
Below we assume $\epsilon \sim N(0, D)$, with D diagonal.

If r is a forecast, and V is a design matrix, we have

$$r = V^T \cdot f + \epsilon$$

The vector f is determined here by regression, and so is the assumed distribution of ϵ.
Below we assume $\epsilon \sim N(0, D)$, with D diagonal.

If r is a forecast, and V is a design matrix, we have

$$r = V^T \cdot f + \epsilon$$

The vector f is determined here by regression, and so is the assumed distribution of ϵ.

Below we assume $\epsilon \sim N(0, D)$, with D diagonal.

If r is a forecast, and V is a design matrix, we have

$$r = V^T \cdot f + \epsilon$$

The vector f is determined here by regression, and so is the assumed distribution of ϵ.
Below we assume $\epsilon \sim N(0, D)$, with D diagonal.

G/I suggest the robust problem based on temporal regression

$$\max_{\{w \in \mathcal{W}\}} \min_{V \in S_v, \mu \in S_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}.$$

with

- $S_v = \left\{V : V = V_0 + W, \|W_i\|_g \leqslant \rho_i, i = 1, ..., n\right\}$
- $S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leqslant \gamma_i, i = 1, ..., n\}$
- $S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \ldots, n\}$

Here, W_i denotes the ith column of W and $\|w\|_g = \sqrt{w^T G w}$

G/I suggest the robust problem based on temporal regression

$$\max_{\{w \in \mathcal{W}\}} \min_{V \in S_v, \mu \in S_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}.$$

with

- $S_v = \left\{V : V = V_0 + W, \|W_i\|_g \leqslant \rho_i, i = 1, ..., n\right\}$
- $S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leqslant \gamma_i, i = 1, ..., n\}$
- $S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \ldots, n\}$

Here, W_i denotes the ith column of W and $\|w\|_g = \sqrt{w^T G w}$

G/I suggest the robust problem based on temporal regression

$$\max_{\{w \in \mathcal{W}\}} \min_{V \in S_v, \mu \in S_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}.$$

with

- $S_v = \left\{V : V = V_0 + W, \|W_i\|_g \leqslant \rho_i, i = 1, ..., n\right\}$
- $S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leqslant \gamma_i, i = 1, ..., n\}$
- $S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \ldots, n\}$

Here, W_i denotes the ith column of W and $\|w\|_g = \sqrt{w^T G w}$

G/I suggest the robust problem based on temporal regression

$$\max_{\{w \in \mathcal{W}\}} \min_{V \in S_v, \mu \in S_m, D \in S_d} \frac{\mu^T w - r_f}{\sqrt{w^T \Sigma w}}.$$

with

- $S_v = \left\{ V : V = V_0 + W, \|W_i\|_g \leqslant \rho_i, i = 1, ..., n \right\}$
- $S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leqslant \gamma_i, i = 1, ..., n\}$
- $S_d = \{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}), d_i \in [\underline{d}_i, \overline{d}_i], i = 1, \ldots, n\}$

Here, W_i denotes the ith column of W and $\|w\|_g = \sqrt{w^T G w}$

As noted earlier, G/I determine their robust model based on regression.

In particular, G, γ, and ρ are based on confidence intervals and (temporal) regression coefficients.

The resulting formula may be written (and solved) as a second order cone program

We alter the construction of G, γ, and ρ as we are interested in cross-sectional regressions.

As noted earlier, G/I determine their robust model based on regression.

In particular, G, γ, and ρ are based on confidence intervals and (temporal) regression coefficients.

The resulting formula may be written (and solved) as a second order cone program

We alter the construction of G, γ, and ρ as we are interested in cross-sectional regressions.

As noted earlier, G/I determine their robust model based on regression.

In particular, G, γ, and ρ are based on confidence intervals and (temporal) regression coefficients.

The resulting formula may be written (and solved) as a second order cone program

We alter the construction of G, γ, and ρ as we are interested in cross-sectional regressions.

As noted earlier, G/I determine their robust model based on regression.

In particular, G, γ, and ρ are based on confidence intervals and (temporal) regression coefficients.

The resulting formula may be written (and solved) as a second order cone program

We alter the construction of G, γ, and ρ as we are interested in cross-sectional regressions.

We consider

$$B = [f_{t-p} \; \cdots \; f_{t-2} \, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T \right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}, \;\; \rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

We consider

$$B = [f_{t-p} \ \cdots \ f_{t-2} \, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T\right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}, \ \rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

We consider

$$B = [f_{t-p} \ \cdots \ f_{t-2}\, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T\right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}$, $\rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

We consider

$$B = [f_{t-p} \ \cdots \ f_{t-2}\, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T\right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}, \ \ \rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

We consider

$$B = [f_{t-p} \ \cdots \ f_{t-2}\, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T\right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}, \ \ \rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

We consider

$$B = [f_{t-p} \ \cdots \ f_{t-2} \, f_{t-1}].$$

$$G = \left(B \cdot B^T - (B \cdot \mathbf{1}) \cdot (B \cdot \mathbf{1})^T \right)$$

and $\gamma_i = \sqrt{(B \cdot B^T)^{-1}_{(1,1)} c_{p,m}(\omega) \cdot \sigma_i^2}, \ \ \rho_i = \sqrt{c_{p,m}(\omega) \cdot \sigma_i^2}$

We have assumed that f(1) is the coefficient related to the intercept of our model

We also modify the expected return vector from their formulation.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.

There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.
There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.

There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.

There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.
There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Using the above uncertainty sets, and a 95% confidence interval, we obtain robust portfolios seeking to maximize the Sharpe ratio.
There are several details to keep in mind when performing backtests, such as,

- awareness of survivorship bias,
- resisting the temptation to snoop the data,
- and a concerted effort must be made to introduce 'future' information.

I took all of this into consideration and attempted to not introduce any of the above failings in the results that follow.

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation
We assume a one month holding period below

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation
We assume a one month holding period below

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation

We assume a one month holding period below

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation

We assume a one month holding period below

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation

We assume a one month holding period below

Below we include results for three problems:

- Maximizing the Sharpe ratio using the trailing mean
- Maximizing the Sharpe ratio using a forecasted return based on the model above
- Maximizing the Sharpe ratio using the robust counterpart cited and the factor model

Our universe consists of large cap stocks - those with market cap over $10 billion at the time of portfolio formation

We assume a one month holding period below

		S&P 500	Nominal Sharpe μ: Trailing Mean	Nominal Sharpe μ: forecast	Robust Sharpe μ: forecast
Annual Return (%)	2002 (beginning March)	-17.15	0.33	-0.34	0.21
	2003	13.02	5.20	4.92	2.10
	2004	10.93	1.03	2.64	11.01
	2005	6.45	24.94	5.62	7.30
	2006	12.10	-10.69	-0.27	4.55
	2007	5.75	57.13	21.78	-0.49
	2008	-39.49	2.49	3.36	8.78
	2009 (through June)	2.56	-26.52	18.38	-1.21
	Annualized Return (%)	-2.78	4.78	7.37	4.31
	Annualized Volatility	15.78	24.28	15.93	6.19
	Annualized Sharpe Ratio	(0.10)	0.32	0.53	0.71
	Max Gain	9.39	18.86	15.85	6.82
	Max Drawdown	-16.94	-26.69	-14.36	-4.85
S&P Relative	β	1.00	(0.11)	(0.09)	(0.06)
	α	0.00	0.63	0.69	0.36

The above table has a few interesting takeaways:

- Addressing the nonstationarity of underlying returns, *even in this simple model*, produces superior results to assuming the past is like the future; viz., compare nominal problems above.
- Addressing model uncertainty provides a significant tool to enhance performance by various metrics; viz., compare nominal with forecast to robust.

The opportunity set for those willing to understand finance through the lens of an empirical mathematician is riper than ever.

The above table has a few interesting takeaways:

- Addressing the nonstationarity of underlying returns, *even in this simple model*, produces superior results to assuming the past is like the future; viz., compare nominal problems above.
- Addressing model uncertainty provides a significant tool to enhance performance by various metrics; viz., compare nominal with forecast to robust.

The opportunity set for those willing to understand finance through the lens of an empirical mathematician is riper than ever.

The above table has a few interesting takeaways:

- Addressing the nonstationarity of underlying returns, *even in this simple model*, produces superior results to assuming the past is like the future; viz., compare nominal problems above.
- Addressing model uncertainty provides a significant tool to enhance performance by various metrics; viz., compare nominal with forecast to robust.

The opportunity set for those willing to understand finance through the lens of an empirical mathematician is riper than ever.

The above table has a few interesting takeaways:

- Addressing the nonstationarity of underlying returns, *even in this simple model*, produces superior results to assuming the past is like the future; viz., compare nominal problems above.
- Addressing model uncertainty provides a significant tool to enhance performance by various metrics; viz., compare nominal with forecast to robust.

The opportunity set for those willing to understand finance through the lens of an empirical mathematician is riper than ever.

I would like to thank the IMA for holding their Mathematical Modeling in Industry Workshops.
Much of the data shown in the above table was calculated by students from the 2009 workshop (XIII).

fin.